

AP20 Rec'd PCT/PTO 17 MAY 2006

## Method of Detecting Mutations Associated with Thrombosis

## BACKGROUND OF THE INVENTION

## FIELD OF THE INVENTION

[0001] The present invention relates to methods and kits for the detection mutations associated with thrombosis.

## DESCRIPTION OF THE PRIOR ART

[0002] Venous thromboembolism (VTE) disease also called venous thrombosis, which manifests clinically as deep vein thrombosis (DVT) and pulmonary embolisms (PE), represents a major health problem worldwide. A DVT is a blood clot that forms in leg veins. A PE is caused by a blood clot, typically in the leg, groin, or pelvic veins (and occasionally in upper extremity veins), which breaks free and travels to the lung arteries. A DVT is often the source of a clot that travels to the lung arteries and becomes a PE.

[0003] VTE is a multifactorial disease which is caused not only by multiple genetic factors but also by multiple acquired or environmental risk factors, such as surgery, use of oral contraceptives, hormone replacement therapy, and advanced age. The incidence of symptomatic venous thrombosis cases is approximately 1 in 1000 people per year. Various gene-gene, gene-environment, and environment-environment interactions between risk factors work synergistically to increase the risk of an individual to VTE. Despite recognition of risk factors and availability of pharmacologically effective options for prophylaxis, DVT and PE remain common causes of morbidity and mortality.

[0004] Several single nucleotide polymorphisms (SNPs) associated with VTE have been identified. The factor V Leiden polymorphism (G1691A) (Genbank Accession #Z99572) has been identified as the most common inherited cause, and is implicated in 20 to 40% of venous thrombosis cases. Heterozygotes possess a 3 to 7-fold increased risk of thrombosis while homozygous mutants carry a 50 to 100-fold increased risk. The second most common cause of inherited thrombophilia involves the factor II (prothrombin) G20210A polymorphism (Genbank Accession #M17262). This mutation results in elevated levels of factor II through mRNA stabilization and accounts for 6 to 8% of venous thrombosis cases. Factor II heterozygotes carry about a 2 to 5-fold increased risk of venous thromboembolism in the absence of other risk factors and a more than additive synergistic risk when other risk factors, especially oral contraceptive use, are present. Approximately 10 to 12% of factor V

5 heterozygotes with venous thrombosis also carry the factor II mutation. A third independent risk factor for thrombosis, is the methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism (Genbank Accession #NM005957.1). This mutation produces an MTHFR protein with reduced activity resulting in an increased serum level of homocysteine. This polymorphism is extremely common with homozygote frequencies of 5 to 15%.  
10 Furthermore, this polymorphism is also a risk factor for atherosclerotic heart disease, pre-eclampsia and fetal neural tube defects. There is a general consensus that the three aforementioned SNPs are the most important mutations associated with thrombophilia. Three additional SNPs, MTHFR A1298C (Genbank Accession #NM005957.1), factor XIII val34leu (factor XIII G4377T Genbank Accession #AF418272) and tissue factor plasma inhibitor  
15 (TFPI) C536T (GenBank Accession #M59497), are believed to have little or no independent effect on venous thrombosis. However, they may act synergistically with other genetic or acquired risk factors resulting in a more than additive effect or, in the case of factor XIII val34leu, a protective effect.

[0005] Individual polymorphisms, as described above, may have little or no independent  
20 effect on venous thrombosis but may act synergistically with other genetic or acquired risk factors, resulting in a more than additive effect. While approaches for providing this information by analyzing one gene at a time are currently available, the ability to detect all of the above mutations simultaneously would be extremely useful in establishing a complete picture of a patient's genetic risk profile.

25 [0006] Methodologies which can be used to detect the above mentioned SNPs are characterized by specific deficiencies. There are currently no rapid methods for determining several mutations across a number of genes simultaneously. For example, DNA sequencing of the above mentioned alleles for a large number of samples would require several days for simplex Polymerase Chain Reaction (PCR) amplification of individual SNPs, amplicon  
30 purification, sequencing reaction set-up and electrophoresis. This would have to be repeated for each of the amplicons generated representing the individual SNPs. Analysis of the sequencing data involves additional time.

#### **Multiplex Allele Specific Primer Extension and Solid Support Detection of SNPs**

[0007] Multiplex allele specific primer extension, and hybridization of extended primers  
35 to a solid support is described generally in the prior art. ASPE technology has been generally described in U.S. Patent No. 4,851,331. The technology is designed to identify the presence or absence of specific polymorphic sites in the genome.

5 [0008] Multiplex ASPE in conjunction with hybridization to a support for mutation detection can be described generally as follows:

[0009] 1) Amplifying regions of DNA comprising polymorphic loci utilizing a multiplexed, PCR.

10 [0010] 2) Allele specific extension of primers wherein the amplified regions of DNA serve as target sequences for the allele specific extension. Extension primers that possess a 3' terminal nucleotide which form a perfect match with the target sequence are extended to form extension products. Modified nucleotides are incorporated into the extension product, such nucleotides effectively labelling the extension products for detection purposes. Alternatively, an extension primer may instead comprise a 3' terminal nucleotide which forms a mismatch  
15 with the target sequence. In this instance, primer extension does not occur unless the polymerase used for extension possesses exonuclease activity.

[0011] 3) Hybridizing the extension product to a probe on a solid support, such as a microarray, wherein the probe is complementary to the 5' end of the extension product.

20 [0012] The extension primers used in a methodology as described above, possess unique sequence tags at their 5' ends. For example, the sequence tags may allow the extension products to be captured on a solid support.

[0013] Variations of the above technology have been described, for example, in U.S. Patent No. 6,287,778 and PCT Application (WO 00/47766).

25 [0014] It is an object of the present invention to provide a method for the detection of variants associated with thrombosis.

## SUMMARY OF THE INVENTION

30 [0015] In one aspect, the present invention provides a method for detecting the presence or absence of a variant nucleotide in at least two SNP sites associated with thrombosis, the SNP sites selected from the group consisting of Factor V Leiden G1691A, Prothrombin (Factor II) G20210A, MTHFR C677T, MTHFR A1298C, Factor XIII G4377T, and tissue factor plasma inhibitor (TFPI) C536T, the method comprising the steps of,

[0016] a) amplifying regions of DNA containing the at least two SNP sites to form amplified DNA products;

35 [0017] b) hybridizing at least two tagged allele specific extension primers to a complementary target sequence in the amplified DNA products, wherein each tagged allele specific extension primer has a 3'-end hybridizing portion substantially complementary to an

5 allele of one of the SNP sites associated with thrombosis and a 5'-end tag portion complementary to one of a set probes, the terminal nucleotide of the 3' end hybridizing portion being either complementary to a suspected variant nucleotide or to the corresponding wild type nucleotide of the SNP site;

[0018] c) extending the at least two tagged allele specific extension primers, using  
10 labelled nucleotides, if the terminal nucleotide of the 3' end hybridizing portion is a perfect match to an allele of one of the SNP sites in the amplified DNA products;

[0019] d) hybridizing the at least two allele two tagged allele specific extension primers to the set of probes and detecting the presence of labelled extension products.

[0020] In another aspect, in the method of the present invention the at least two tagged  
15 allele specific extension primers have a 3' end hybridizing portion corresponding to bases 25 and up of any two of SEQ ID NO: 1 to SEQ ID NO: 12.

[0021] In another aspect, in the method of the present invention the at least two tagged allele specific extension primers have a 5'-end tag portion comprising a sequence corresponding to bases 1 to 24 of any two of SEQ ID NO: 1 to SEQ ID NO: 12.

20 [0022] In another aspect, in the method of the present invention the at least two tagged allele-specific extension primers are selected from the group consisting of SEQ ID NO: 1 to SEQ ID NO: 12.

[0023] In another aspect, the present invention provides a kit for use in detecting the presence or absence of a variant nucleotide in at least two SNP sites associated with  
25 thrombosis, the SNP sites selected from the group consisting of Factor V Leiden G1691A, Prothrombin (Factor II) G20210A, MTHFR C677T, MTHFR A1298C, Factor XIII G4377T, and tissue factor plasma inhibitor (TFPI) C536T, the kit comprising a set of at least two tagged allele specific extension primers wherein each tagged allele specific extension primer has a 3'-end hybridizing portion substantially complementary to a first allele of one of the  
30 SNP sites associated with thrombosis and a 5'-end tag portion complementary to one of a set probes.

[0024] In another aspect, the present invention provides a kit for use in detecting the presence or absence of a variant nucleotide in at least two SNP sites associated with thrombosis, the SNP sites selected from the group consisting of Factor V Leiden G1691A,  
35 Prothrombin (Factor II) G20210A, MTHFR C677T, MTHFR A1298C, Factor XIII G4377T, and tissue factor plasma inhibitor (TFPI) C536T, the kit including at least two tagged allele

5 specific extension primers having a 3' end hybridizing portion corresponding to bases 25 and up of any two of SEQ ID NO: 1 to SEQ ID NO: 12.

[0025] In another aspect, the present invention provides a kit for use in detecting the presence or absence of a variant nucleotide in at least two SNP sites associated with thrombosis, the SNP sites selected from the group consisting of Factor V Leiden G1691A, Prothrombin (Factor II) G20210A, MTHFR C677T, MTHFR A1298C, Factor XIII G4377T, and tissue factor plasma inhibitor (TFPI) C536T, the kit including at least two tagged allele specific extension primers selected from the group consisting of SEQ ID NO: 1 to SEQ ID NO: 12.

[0026] In another aspect, the present invention provides a kit for use in detecting the presence or absence of a variant nucleotide in at least two SNP sites associated with thrombosis, the SNP sites selected from the group consisting of Factor V Leiden G1691A, Prothrombin (Factor II) G20210A, MTHFR C677T, MTHFR A1298C, Factor XIII G4377T, and tissue factor plasma inhibitor (TFPI) C536T, the kit including at least two tagged allele specific primers having a 5'-end tag portion comprising a sequence corresponding to bases 1 to 24 of any two of SEQ ID NO: 1 to SEQ ID NO: 12.

[0027] In another aspect the present invention provides a kit for use in detecting the presence or absence of a variant nucleotide in at least two SNP sites associated with thrombosis, the SNP sites selected from the group consisting of Factor V Leiden G1691A, Prothrombin (Factor II) G20210A, MTHFR C677T, MTHFR A1298C, Factor XIII G4377T, and tissue factor plasma inhibitor (TFPI) C536T, the kit comprising a set of PCR amplification primers for amplifying regions of DNA containing the at least two SNP sites, the set comprising at least two pairs of PCR primers selected from the group of pairs consisting of:

[0028] SEQ ID NO: 13 and SEQ ID NO: 14, SEQ ID NO: 15 and SEQ ID NO: 16, SEQ ID NO: 17 and SEQ ID NO: 18, SEQ ID NO: 19 and SEQ ID NO: 20, SEQ ID NO: 21 and SEQ ID NO: 22, and SEQ ID NO: 23 and SEQ ID NO: 24.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

[0029] These and other features of the preferred embodiments of the invention will become more apparent in the following detailed description in which reference is made to the appended drawings wherein:

[0030] Figure 1 depicts a general overview of steps of the present invention.

5 [0031] Figure 2 depicts the high specificity of PCR primers used for multiplex amplification of SNPs associated with thrombosis.

[0032] Figure 3 depicts the high specificity of ASPE primers for specific alleles.

[0033] Figure 4 graphically depicts the ability of the present invention to distinguish between homozygous wild type, heterozygous, and homozygous mutant individuals.

10 [0034] Figure 5 depicts an overview of methodologies used to detect polymorphisms associated with thrombosis.

[0035] Figure 6 presents results for genotyping using a Luminex Bead-Based platform.

[0036] Figure 7 presents results for genotyping using a Metrigenix Flow-Thru Chip platform.

## 15 DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0037] **Definitions:**

[0038] The following terms used in the present application will be understood to have the meanings defined below:

[0039] The term single nucleotide polymorphism (SNP), as used herein represents one of  
20 the most common forms of genetic variation. These polymorphisms occur when a single nucleotide (A, G, C or T) in the genome is altered. SNPs generally tend to be evolutionarily stable from generation to generation and, as such, can be used to study specific genetic abnormalities throughout a population. SNPs often occur in protein coding regions and, as a result, may lead to the expression of a defective or variant form of a protein. Such  
25 polymorphisms can therefore serve as effective indicators of genetic disease. However, not all SNPs are found in protein coding regions of the genome. Some SNPs are located in noncoding regions, but these polymorphisms may also lead to altered protein expression. Specifically, SNP sites in noncoding regions may, for example, lead to differential and defective splicing. In diseases where multiple genes may influence the onset of the disease,  
30 SNPs can be used as diagnostic tools for identifying individuals with a predisposition for manifesting the disease, genotyping the patients suffering from the disease in terms of the genetic causes underlying the condition, and facilitating drug development based on the insight revealed regarding the role of target proteins in the pathogenesis process.

[0040] The terms "oligonucleotide" and "polynucleotide" as used in the present  
35 application refer to DNA sequences being of greater than one nucleotide in length. Such sequences may exist in either single or double-stranded form. Examples of oligonucleotides described herein include PCR primers, ASPE primers, and anti-tags.

5 [0041] The term “allele” is used herein to refer to variants of a nucleotide sequence.

[0042] The expression “allele specific primer extension (ASPE)”, as used herein, refers to a mutation detection method utilizing primers which hybridize to a corresponding DNA sequence and which are extended depending on the successful hybridization of the 3’ terminal nucleotide of such primer. Amplified regions of DNA serve as target sequences for

10 ASPE primers. Extension primers that possess a 3’ terminal nucleotide which form a perfect match with the target sequence are extended to form extension products. Modified nucleotides can be incorporated into the extension product, such nucleotides effectively labelling the extension products for detection purposes. Alternatively, an extension primer may instead comprise a 3’ terminal nucleotide which forms a mismatch with the target

15 sequence. In this instance, primer extension does not occur unless the polymerase used for extension inadvertently possesses exonuclease activity.

[0043] The term “genotype” refers to the genetic constitution of an organism. More specifically, the term refers to the identity of alleles present in an individual. “Genotyping” of an individual or a DNA sample refers to identifying the nature, in terms of nucleotide base,

20 of the two alleles possessed by an individual at a known polymorphic site.

[0044] The term “polymorphism”, as used herein, refers to the coexistence of more than one form of a gene or portion thereof.

[0045] The term “PCR”, as used herein, refers to the polymerase chain reaction. PCR is a method of amplifying a DNA base sequence using a heat stable polymerase and a pair of

25 primers, one primer complementary to the (+)-strand at one end of the sequence to be amplified and the other primer complementary to the (-) strand at the other end of the sequence to be amplified. Newly synthesized DNA strands can subsequently serve as templates for the same primer sequences and successive rounds of heat denaturation, primer annealing and strand elongation results in rapid and highly specific amplification of the

30 desired sequence. PCR can be used to detect the existence of a defined sequence in a DNA sample.

[0046] The term “primer”, as used herein, refers to a short single-stranded oligonucleotide capable of hybridizing to a complementary sequence in a DNA sample. A primer serves as an initiation point for template dependent DNA synthesis.

35 Deoxyribonucleotides can be joined to a primer by a DNA polymerase. A “primer pair” or “primer set” refers to a set of primers including a 5’upstream primer that hybridizes with the complement of the 5’ end of the DNA sequence to be amplified and a 3’ downstream primer

5 that hybridizes with the 3' end of the DNA sequence to be amplified. The term "PCR primer" as used herein refers to a primer used for a PCR reaction. The term "ASPE primer" as used herein refers to a primer used for an ASPE reaction.

[0047] The term "tag" as used herein refers to an oligonucleotide sequence that is coupled to an ASPE primer. The sequence is generally unique and non-complementary to the human genome while being substantially complementary to a probe sequence. The probe sequence may be, for example, attached to a solid support. Tags serve to bind the ASPE primers to a probe.

[0048] The term "tagged ASPE primer" as used herein refers to an ASPE primer that is coupled to a tag.

15 [0049] The term "anti-tag" or "probe" as used herein refers to an oligonucleotide sequence having a sequence complementary to, and capable of hybridizing to, the tag sequence of an ASPE primer. The "anti-tag" may be coupled to a support.

[0050] The term "wild type" as used herein refers to the normal, or non-mutated, or functional form of a gene.

20 [0051] The term "homozygous wild-type" as used herein refers to an individual possessing two copies of the same allele, such allele characterized as being the normal and functional form of a gene.

[0052] The term "heterozygous" as used herein refers to an individual possessing two different alleles of the same gene.

25 [0053] The term "homozygous mutant" as used herein refers to an individual possessing two copies of the same allele, such allele characterized as the mutant form of a gene.

[0054] The term "mutant" as used herein refers to a mutated, or potentially non-functional form of a gene.

#### **DESCRIPTION OF THE INVENTION**

30 [0055] The present invention was developed in response to a need for a rapid, highly specific, and cost-effective method to simultaneously identify multiple genetic risk factors associated with thrombosis. Such identification of risk factors can enhance both treatment and prevention of serious health problems associated with the disease.

[0056] The present invention provides a novel, multiplex method of detecting multiple SNPs associated with thrombosis. Specifically, the methodology can be used for the detection of the presence of absence of two or more mutations selected from the group consisting of Factor V Leiden G1691A, Prothrombin (Factor II) G20210A, MTHFR C677T,



5 MTHFR A1298C, factor XIII val34leu, and tissue factor plasma inhibitor (TFPI) C536T. The positive detection of two or more of the mutated forms of each of the above polymorphic sites may be indicative of an individual having a predisposition for thrombosis.

[0057] The present invention is further characterized by a high level of specificity. Such specificity is required in order to ensure that any result generated is a true representation of  
10 the genomic target and not simply the result of non-specific interactions occurring between reagents present in reactions. This is especially important for multiplexed DNA-based tests where the numerous sequences present in the reaction mixture, most of which are non-complementary, may interact non-specifically depending on the reaction conditions. The ASPE primer and PCR primer sequences described below have been selected due to their  
15 minimal cross-reactivity.

[0058] The present invention is also characterized by its high level of accuracy when compared to existing methods of detection of SNPs associated with thrombosis, for example, DNA sequencing. Results comparing the two methodologies are provided by example further below.

20 [0059] The methodology of the present invention utilizes the combination of multiplex ASPE technology with hybridization of tagged and labelled extension products to probes in order to facilitate detection. Such methodology is suitable for high-throughput clinical genotyping applications.

[0060] In one aspect, the present invention provides a method for detecting the presence  
25 or absence of a variant nucleotide in at least two SNP sites in a sample selected from the group consisting of Factor V Leiden G1691A, Prothrombin (Factor II) G20210A, MTHFR C677T, MTHFR A1298C, Factor XIII G4377T, and tissue factor plasma inhibitor (TFPI) C536T the method comprising the steps of:

[0061] Amplifying regions of DNA which may contain the above mentioned variants.

30 [0062] Hybridizing at least two tagged allele specific extension primers to a complementary region of amplified DNA, each tagged allele specific primer having a 3' portion complementary to a region of the amplified DNA, a 3' terminal nucleotide complementary to one allele of one of the mutation sites (wild type or mutant) mentioned above, and a 5' portion complementary to a probe sequence.

35 [0063] Extending tagged ASPE primers, whereby a labelled extension product of the primer is synthesised when the 3' terminal nucleotide of the primer is complementary to a corresponding nucleotide in the target sequence; no extension product is synthesised when

5 the terminal nucleotide of the primer is not complementary to the corresponding nucleotide in the target sequence.

[0064] Hybridizing extension products to a probe and detecting labelled extension products. Detection of a labelled extension product is indicative of the presence of the allele complementary to the 3'-terminal nucleotide of the ASPE primer. In the absence of a  
10 labelled extension product, it is determined that the allele corresponding to the 3' end of the ASPE primer is not present in the sample.

[0065] A general overview of one example of the above-mentioned method is presented in figure 1. The present invention should not be limited to the example provided in figure 1. A DNA sample is first prepared 10 using methods known in the art. Multiplex PCR  
15 amplification 20 is conducted in order amplify regions of DNA containing SNP sites that are associated with thrombosis. A multiplex ASPE reaction 30 is then conducted. By example only, 33 illustrates a wild type and a mutant allele of a gene. At step 36 ASPE primers are hybridized to amplified regions of DNA. If the 3' terminal nucleotide of an ASPE primer is complementary to a corresponding nucleotide in the target sequence, a labelled extension  
20 product is formed 39 as will be described further below. The ASPE may be sorted on an addressable universal sorting array 40 wherein the presence of a labelled extension product may be detected using, for example, xMAP detection 50.

#### **DNA Sample Preparation**

[0066] Patient samples can be extracted with a variety of methods known in the art to  
25 provide nucleic acid (most preferably genomic DNA) for use in the following method. In a preferred embodiment, a DNA sample is extracted from whole blood.

#### **Amplification**

[0067] In a first step at least two regions of DNA containing SNP sites associated with thrombosis are amplified.

30 [0068] In a preferred embodiment of the present invention, PCR amplification of regions containing polymorphic sites associated with thrombosis is initiated using at least two pairs of PCR primers selected from the group of primer pairs consisting of: SEQ ID NO: 13 and SEQ ID NO: 14, SEQ ID NO: 15 and SEQ ID NO: 16, SEQ ID NO: 17 and SEQ ID NO: 18, SEQ ID NO: 19 and SEQ ID NO: 20, SEQ ID NO: 21 and SEQ ID NO: 22, and SEQ ID NO:  
35 23 and SEQ ID NO: 24. The relationships of each pair of primers to the polymorphic sites mentioned further above is presented in Table 1.

5 **TABLE 1: Primer Pairs Used to Amplify Regions Containing Thrombosis Associated Polymorphic Sites**

Polymorphic Site	Primer Pair
Factor V Leiden G1691A	SEQ ID NO.: 13 (forward primer) SEQ ID NO.: 14 (reverse primer)
Prothrombin (Factor II) G20210A	SEQ ID NO.: 15 (forward primer) SEQ ID NO.: 16 (reverse primer)
MTHFR C677T	SEQ ID NO.: 17 (forward primer) SEQ ID NO.: 18 (reverse primer)
MTHFR A1298C	SEQ ID NO.: 19 (forward primer) SEQ ID NO.: 20 (reverse primer)
Tissue Factor Plasma Inhibitor (TFPI) C536T	SEQ ID NO.: 21 (forward primer) SEQ ID NO.: 22 (reverse primer)
Factor XIII G4377T	SEQ ID NO.: 23 (forward primer) SEQ ID NO.: 24 (reverse primer)

10 [0069] An individual skilled in the art will recognize that alternate PCR primers could be used to amplify the target polymorphic regions, however, in a preferred embodiment the primers listed in Table 1 are selected due to their minimal non-specific interaction with other sequences in the reaction mixture.

#### ASPE

15 [0070] The ASPE step of the method of the present invention is conducted using at least two tagged ASPE primers selected from the group of ASPE primers consisting of bases 25 and up of SEQ ID NO: 1 to SEQ ID NO: 12.

[0071] The ASPE primer set of the present invention has been optimized, as described further below by example, to ensure high specificity and accuracy of diagnostic tests utilizing such allele specific primers.

[0072] The ASPE primers of the present invention are as described in table 2.

Allele Detected	Allele Specific Primer
Factor V Leiden G1691	SEQ ID NO: 1 (bases 25 to 46)
Factor V Leiden 1691A	SEQ ID NO: 2 (bases 25 to 46)
Prothrombin (Factor II) G20210	SEQ ID NO: 3 (bases 25 to 45)
Prothrombin (Factor II) 20210A	SEQ ID NO: 4 (bases 25 to 45)

MTHFR C677	SEQ ID NO: 5 (bases 25 to 44)
MTHFR 677T	SEQ ID NO: 6 (bases 25 to 44)
MTHFR A1298	SEQ ID NO: 7 (bases 25 to 46)
MTHFR 1298C	SEQ ID NO: 8 (bases 25 to 46)
Tissue Factor Plasma Inhibitor (TFPI) C536	SEQ ID NO: 9 (bases 25 to 43)
Tissue Factor Plasma Inhibitor (TFPI) 536T	SEQ ID NO: 10 (bases 25 to 43)
Factor XIII G4377	SEQ ID NO: 11 (bases 25 to 42)
Factor XIII 4377T	SEQ ID NO: 12 (bases 25 to 42)

5 [0073] The 3' end hybridizing portion of the extension primer is hybridized to the amplified material. Where the 3' terminal nucleotide of an ASPE primer is complementary to the polymorphic site, primer extension is carried out using a modified nucleotide. Where the 3' terminal nucleotide of the ASPE primer is not complementary to the polymorphic region, no primer extension occurs.

10 [0074] In one embodiment, labelling of the extension products is accomplished through the incorporation of biotinylated nucleotides into the extension product which may be identified using fluorescent (Streptavidin-Phycoerythrin) or chemiluminescent (Streptavidin-Horseradish Peroxidase) reactions. However, an individual skilled in the art will recognize that other labelling techniques may be utilized. Examples of labels useful for detection  
15 include but are not limited to radiolabels, fluorescent labels (e.g fluorescein and rhodamine), nuclear magnetic resonance active labels, positron emitting isotopes detectable by a positron emission tomography ("PET") scanner, and chemilumescers such as luciferin, and enzymatic markers such as peroxidase or phosphatase.

[0075] Each ASPE primer used in the methodology as described above, possess a unique  
20 sequence tag at their 5' ends. The sequence tags allow extension products to be detected with a high degree of specificity, for example, through capture on a solid support in order to facilitate detection.

[0076] In a preferred embodiment, tags are selected from the group of tags consisting of any one of SEQ ID NO: 1 to SEQ ID NO: 12 (bases 1 to 24).

25 [0077] **Detection**

[0078] The tagged 5' portions of the allele specific primers of the present invention are complementary to probe sequences. Upon hybridization of the allele specific primers to a corresponding probe sequence the presence of extension products can be detected.

5 [0079] In a preferred embodiment, probes used in the methodology of the present invention are coupled to a solid support, for example a 'universal' bead-based microarray.

[0080] Examples of supports that can be used in the present invention include, but are not limited to, bead based microarrays and 2D glass microarrays. The preparation, use, and analysis of microarrays are well known to persons skilled in the art. (See, for example,  
10 Brennan, T. M. et al. (1995) U.S. Pat. No. 5,474,796; Schena, et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995), PCT Application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R. A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M. J. et al. (1997) U.S. Pat. No. 5,605,662.). Detection can be achieved through arrays using, for example, chemiluminescence or fluorescence technology  
15 for identifying the presence of the SNPs.

[0081] Universal arrays function as sorting tools indirectly detecting the target of interest and are designed to be isothermal and minimally cross-hybridizing as a set. Examples of microarrays which can be used in the present invention include, but should not be limited to, Luminex's® bead based microarray systems, and Metrigenix's™ Flow Thru chip technology.

20 [0082] In one embodiment, for example, Luminex's 100 xMAP™ fluorescence based solid support microarray system is utilized. Anti-tag sequences complementary to the tag regions of the ASPE primers/extension products, described above, are coupled to the surface of internally fluorochrome-color-coded microspheres. An array of anti-tag microspheres is produced, each set of microspheres having its own characteristic spectral address. The  
25 mixture of tagged, extended biotinylated ASPE primers is combined with the array of anti tagged microspheres and allowed to hybridize under stringent conditions.

[0083] In a reaction mixture, a fluorescent reporter molecule (e.g. streptavidin-phycoerythrin) is used to detect labelled extension products which are synthesized when the terminal nucleotide of an ASPE primer is complementary to a corresponding nucleotide in  
30 the target sequence.

[0084] The reaction mixture, comprising microspheres, extension products etc. is injected into a reading instrument, for example Luminex's 100 xMAP™, which uses microfluidics to align the microspheres in single file. Lasers are used to illuminate the colors both internal to the microspheres, and attached to the surface in the form of extension products hybridized to  
35 anti-tag sequences. The Luminex 100 xMAP™, interprets the signal received and identifies the presence of wild type and/or mutant alleles. The presence of the mutant allele of any one or more of the bi-allelic sites consisting of Factor V Leiden G1691A, Prothrombin (Factor II)

5 G20210A, MTHFR C677T, MTHFR A1298C, and tissue factor plasma inhibitor (TFPI) C536T may be indicative of thrombosis, or a pre-disposition to thrombosis. The presence of the mutant allele for factor XIII val34leu has been indicated in providing a weak protective effect against venous thrombosis. Software can be provided which is designed to analyze data associated with the specific extension products and anti-tagged microspheres of the present invention.

[0085] In another embodiment, the Metrigenix Flow-Thru three dimensional microchannel biochip (Cheek, B.J., Steel A.B., Torres, M.P., Yu, Y., and Yang H. Anal. Chem. 2001, 73, 5777-5783) is utilized for genotyping as known in the art. In this embodiment, each set of microchannels represents a different universal anti-tag population.

15 Anti-tag sequences corresponding to the tag regions of the ASPE primers/extension products, described above, are attached to the inner surface of multiple microchannels comprising a cell. Multiple cells make up a chip. The reaction mixture, including biotinylated extension products flows through the cells in the presence of a chemiluminescent reporter substrate such as streptavidin-horseradish peroxidase. Microarray chips can be imaged using technology known in the art, such as an ORCA-ER CCD (Hamamatsu Photonics K. K., Hamamatsu City, Japan), and imaging software, in order to identify the genotype of an individual. The presence of the mutant allele of any one or more of the bi-allelic sites consisting of Factor V Leiden G1691A, Prothrombin (Factor II) G20210A, MTHFR C677T, MTHFR A1298C, and tissue factor plasma inhibitor (TFPI) C536T may be indicative of thrombosis, or a pre-disposition to thrombosis. The presence of the mutant allele for factor XIII val34leu has been indicated in providing a weak protective effect against venous thrombosis. Software can be provided which is designed to analyze data associated with the specific extension products and anti-tagged microspheres of the present invention.

25 [0086] An example of the use of the above-mentioned platforms, which also illustrates the specificity and accuracy of the method of the present invention, is provided below.

#### Kits

[0087] In an additional aspect, the present invention provides kits for the multiplex detection of SNPs associated with thrombosis.

35 [0088] A kit according to the present invention may contain the following components including: a PCR primer mix for amplifying regions containing SNP sites of interest optionally including dNTPs), an ASPE primer mix for generation of labelled extension products (optionally including dNTPs) and a solid support, such as microarray beads, the

5 support having anti-tags complementary to the tagged regions of the ASPE primers. In addition, an individual skilled in the art would recognize other components which could be included in such kits including, for example, buffers and polymerases.

[0089] Kits of the present invention may include PCR primer pairs, ASPE primers, and tagged supports for all six of the SNP sites or may be customized to best suit the needs of an individual end user. For example, if a customer wishes to detect only three of the SNP sites associated with thrombosis, a kit can be customized to include only the PCR primer pairs, ASPE primers, and support required for the detection of the desired SNP sites. As such, the end user of the product can design a kit to match their requirements. In addition, the end user can also control the tests to be conducted at the software level when using, for example, a universal bead based-microarray for detection. For example, software can be provided with a kit, such software reading only the beads for the desired SNPs or by reporting only the results from the desired SNP data. Similar control of data reporting by software can be obtained when the assay is performed on alternate platforms.

[0090] Although the present method has been described in relation to six specific thrombosis associated SNP sites, PCR primers and ASPE primers used to detect additional SNP sites could be included in the above methodology and kits.

[0091] All publications, patents and patent applications are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety

[0092] The examples presented below are provided to illustrate the present invention and are not meant to limit the scope of the invention as will be apparent to persons skilled in the art.

[0093] **EXAMPLE #1: ASPE/Microarray Detection of SNPs Associated with Thrombosis**

[0094] **MATERIALS and METHODS**

[0095] **1) Oligonucleotides**

[0096] All oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA). PCR primers were unmodified and were purified by standard desalting procedures. Universal anti-tags (probes) were 3'-C7 amino-modified for coupling to carboxylated microspheres. All anti-tags were reverse phase HPLC-purified. Chimeric ASPE primers which consisted of a 24mer universal tag sequence 5' to the allele-specific

5 (18-22mer) sequence were also unmodified but were purified by polyacrylamide gel electrophoresis. Following reconstitution, exact oligo concentrations were determined spectrophotometrically using extinction coefficients provided by the supplier. Reconstituted oligos were scanned between 200 and 800 nm and absorbance was measured at 260 nm to calculate oligo concentration.

10 **[0097] 2) Reagents**  
**[0098]** Platinum Taq, Platinum Tsp, individual dNTPs and biotin-dCTP were purchased from Invitrogen Corporation (Carlsbad, CA). Shrimp alkaline phosphatase and exonuclease I were purchased from USB Corporation (Cleveland, OH). Carboxylated fluorescent microspheres were provided by Luminex Corporation (Austin, TX). The EDC cross-linker  
15 (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride) was purchased from Pierce (Rockford, IL). OmniPur reagents including MES (2-(N-morpholino)ethane sulfonic acid), 10% SDS, NaCl, Tris, Triton X-100, Tween-20 and TE buffer were purchased from EM Science (Darmstadt, Germany). The streptavidin-conjugated phycoerythrin was obtained from Molecular Probes Inc. (Eugene, OR).

20 **[0099] 3) Patient Samples**  
**[00100]** Genomic DNA from 132 study participants was isolated from whole blood using the Qiagen QIAamp Blood Kit (Qiagen). DNA samples were stored at -20°C upon isolation. Following shipment, samples were quantified spectrophotometrically by measuring absorbance at 260 nm. Genomic DNA samples were then diluted to 5 ng/uL and stored at  
25 4°C.

**[00101] 4) Genotyping**  
**[00102]** a) MULTIPLEX PCR (6-plex): Multiplex PCR was carried out using 25 ng genomic DNA in a 25 uL final volume. A 'no target' PCR negative control was included with each assay run. The reaction consisted of 30 mmol/L Tris-HCl, pH 8.4, 75 mmol/L  
30 KCl, 2 mmol/L MgCl<sub>2</sub>, 200 umol/L each dNTP, 1.25 units Platinum Taq, with primers ranging from 0.2 to 0.5 umol/L. Samples were cycled in an MJ Research PTC-200 thermocycler (Reno, NV) with cycling parameters set at 95°C for 5 minutes followed by 30 cycles at 95°C for 30 seconds, 58°C for 30 seconds and 72°C for 30 seconds. Samples were then held at 72°C for 5 minutes and kept at 4°C until use.

35 **[00103]** b) ALLELE-SPECIFIC PRIMER EXTENSION: Prior to the ASPE reaction, each PCR reaction was treated with shrimp alkaline phosphatase (SAP) to inactivate any remaining nucleotides (particularly dCTP) so that biotin-dCTP could be efficiently



5 incorporated during the primer extension reaction. Each PCR reaction was also treated with  
exonuclease I (EXO) to degrade remaining PCR primers in order to avoid any interference  
with the tagged ASPE primers and the extension reaction itself. To each 25 uL PCR reaction,  
2 uL SAP (= 2 units) and 0.5 uL EXO (= 5 units) were added directly. Samples were then  
10 incubated at 37°C for 30 minutes followed by a 15 minute incubation at 99°C to inactivate  
the enzymes. Samples were then added directly to the ASPE reaction.

[00104] Multiplex ASPE was carried out using 5 uL of treated PCR product in a final  
volume of 20 uL. Each reaction consisted of 20 mmol/L Tris-HCl pH 8.4, 50 mmol/L KCl,  
1.25 mmol/L MgCl<sub>2</sub>, 5 umol/L biotin-dCTP, 5 umol/L each of dATP, dGTP and dTTP, 1.5  
units Platinum Tsp and 25 nmol/L ASPE primer pool (ie. each ASPE primer present at 500  
15 fmol/reaction). The ASPE reactions were incubated at 96°C for 2 minutes and then subjected  
to 40 cycles at 94°C for 30 seconds, 54°C for 30 seconds and 74°C for 60 seconds. Reactions  
were then held at 4°C until use.

[00105] c) BEAD COUPLING: Amino-modified anti-tag sequences were coupled to  
carboxylated microspheres following Luminex's one-step carbodiimide coupling procedure.  
20 Briefly,  $5 \times 10^6$  microspheres were combined with 1 nmol NH<sub>2</sub>-oligo in a final volume of 50  
uL 0.1 mol/L MES, pH 4.5. A 10 mg/mL EDC working solution was prepared just prior to  
use and 2.5 uL was added to the bead mixture and incubated for 30 minutes. A second 2.5 uL  
aliquot of freshly prepared EDC was added followed by an additional 30 minute incubation.  
Following washes in 0.02% (v/v) Tween-20 and 0.1% (w/v) SDS, the anti-tag coupled beads  
25 were resuspended in 100 uL TE buffer (10 mmol/L Tris, pH 8.0, 1 mmol/L EDTA). Bead  
concentrations were determined using a Beckman Coulter Z2 Particle Count and Size  
Analyzer (Coulter Corp, Miami FL).

[00106] d) UNIVERSAL ARRAY HYBRIDIZATION: Each hybridization reaction was  
carried out using approximately 2500 beads of each of the twelve anti-tag bearing bead  
30 populations. The beads were combined in hybridization buffer (0.22 mol/L NaCl, 0.11 mol/L  
Tris, pH 8.0 and 0.088% (v/v) Triton X-100) and 45 uL of the mix were added to each well of  
an MJ Research 96-well plate (Reno, NV). A 5 uL aliquot of each ASPE reaction was then  
added directly to each well. The samples were then heated to 96°C for 2 minutes in an MJ  
Research PTC-200 followed by a one hour incubation at 37°C. Following this incubation,  
35 samples were filtered through a 1.2 um Durapore Membrane (Millipore Corp, Bedford, MA)  
and washed once using wash buffer (0.2 mol/L NaCl, 0.1 mol/L Tris, pH 8.0 and 0.08% (v/v)  
Triton X-100). The beads were then resuspended in 150 uL reporter solution (1 ug/mL

streptavidin-conjugated phycoerythrin in wash buffer) and incubated for 15 minutes at room temperature. The reactions were read on the Luminex xMAP. Acquisition parameters were set to measure 100 events per bead population and a 100 uL sample volume. A gate setting was established prior to running the samples and maintained throughout the course of the study.

**[00107] e) GENOTYPING BY DNA SEQUENCING**

**[00108]** For all samples used in the study, genotyping results obtained using the method of the present invention were compared to genotyping results obtained using dideoxy dye-terminator sequencing chemistry. For each of the 6 SNPs, sequencing was performed in both the forward and reverse directions on PCR amplimers obtained using individual primer pairs.

**[00109] DATA ANALYSIS AND INTERPRETATION**

**[00110] a) DNA Sequencing:** For all 132 patient samples in the study, forward and reverse DNA sequencing data for each of the 6 SNPs within a sample was analyzed. For a sequencing call to be accepted for a particular SNP, a consensus between the two individuals was required. If a consensus was not obtained for a particular SNP, the call was determined to be ambiguous and was eliminated from the study. In addition, if the forward and reverse sequencing data for a particular SNP did not coincide, the call was also determined to be ambiguous and was eliminated from the accuracy study.

**[00111] b) Genotyping Assay:** For each DNA sample tested using the method of the present invention, median fluorescent intensity (MFI) units were collected for each of the 12 bead populations corresponding to each allele within the assay. For each allele, the NET MFI was calculated by subtracting the 'no target' (PCR negative control) MFI values from the MFI values for each allele of a given sample. In order to exclude samples containing insufficient or degraded DNA or samples generating sub-optimal results, acceptance criteria were defined such that, for each allele within the assay, MFI units were required to be at least 10x the 'no target' MFI for that allele AND at least 300. If these criteria were not met, the SNP was not called and was excluded from the accuracy study. For SNPs meeting the data requirements, the genotype was then determined based on the mutant allelic ratio where:

$$\text{[00112] Mutant Allelic Ratio} = \frac{(\text{NET MFI})_{\text{mutant allele}}}{(\text{NET MFI})_{\text{mutant allele}} + (\text{NET MFI})_{\text{wild-type allele}}}$$

**[00113]** In other words, the mutant allelic ratio represents the fraction of the total net MFI signal for a given SNP attributed to the presence of the mutant allele. By setting cut-off values, the allelic ratio is used to discriminate homozygous wild-type, homozygous mutant

5 and heterozygous SNP calls. Cut-off values were empirically determined for each individual SNP. Typically, the mutant allelic ratio ranged from 0.00 to 0.10 for homozygous wild-type calls, 0.30 to 0.70 for heterozygous calls and 0.90 to 1.00 for homozygous mutant calls.

## **[00114] RESULTS**

### **[00115] a) Assay Optimization**

10 **[00116]** The PCR and ASPE components of the genotyping methodology of the present invention were optimized to ensure high specificity, in addition to sufficient yield, when assessed by PCR amplimer intensity and MFI units generated by the Luminex xMAP system.

**[00117]** For optimal PCR, buffer composition, cycling parameters, annealing temperature, genomic DNA input as well as primer concentrations for each SNP were examined. PCR  
15 products generated under the final optimized conditions were analyzed by gel electrophoresis using the Helixx SuperGel150 system (Scarborough, ON) which is capable of resolving single basepair differences within products. A gel image of 6 patient samples amplified under optimal conditions is given in Figure 2 and clearly demonstrates that the multiplex PCR reaction of the present invention was highly specific for the six desired amplimers.

20 **[00118]** The ASPE component of the method of the present invention was optimized for several parameters affecting specificity and signal output. Parameters examined included cycling parameters, annealing temperature, ASPE primer concentrations and PCR reaction volume added to the ASPE reaction.

**[00119]** Specificity was for the ASPE reaction where non-specific interactions, detrimental  
25 to the end result, may occur between (a) the universal tag sequence and the target alleles and/or (b) the allele-specific primer region and the target alleles. To address this issue of specificity, each allele comprising the thrombophilia panel was PCR-amplified individually and then subjected to ASPE with all twelve universally-tagged ASPE primers present. The ASPE reactions were then sorted on the universal array (ie. all twelve anti-tags present). The  
30 signals generated for each individually-amplified allele in the presence of all twelve genotyping primers shown in Figure 3 clearly indicate that, in the presence of all genotyping primers, each allele reacts only with its corresponding genotyping primer. Two levels of discrimination are built into primer extension genotyping chemistry with one occurring at the hybridization level and the other at the enzymatic extension of a mismatched primer.

35 **[00120]** From the above experiment, it was also evident that the universal anti-tag sequences hybridize specifically to their complements (tags). Extensive studies on the interactions between the 100 anti-tags from which this subset of twelve was derived have

5 been conducted and the performance of the universal tags has been well-characterized under  
varying assay conditions. Multiplexed anti-tag mixes using all 100 bead populations were  
hybridized with each individual complementary tag at various concentrations to check for  
potential cross-hybridization between universal sequences. Significant cross-hybridization  
was not observed, with all data clearly supporting the high specificity built into the design of  
10 the universal tag sequence set.

**[00121] b) Patient Study**

**[00122]** To assess the clinical accuracy of the method of the present invention, a total of  
132 patient DNA samples were genotyped by both the method of the present invention, and  
DNA dideoxy sequencing for each of the following 6 SNPs: factor V Leiden G1691A, factor  
15 II G20210A (prothrombin), MTHFR C677T, MTHFR A1298C, Factor XIII val34leu and  
TFPI C536T. The DNA samples were selected from consecutive patients with venous  
thrombosis or age, sex and ethnicity-matched friend controls.

**[00123]** Overall, the method of the present invention was able to generate calls for 779 out  
of the 792 SNP calls possible for the 132 patient samples analyzed for each of the 6 SNPs. A  
20 total of 13 calls resulting from 3 individual samples could not be made due to minimum  
signal requirements or allelic ratio cut-offs not being met.

**[00124]** The 43 calls not made by sequencing were ambiguous (as defined previously)  
and were divided among 38 samples. To assess the accuracy within the study, the calls which  
could not be made either by the method of the present invention or by sequencing were  
25 excluded from the calculation. A total of 56 SNPs were eliminated from the study for the  
purposes of determining accuracy. Of the remaining 736 SNP calls, 14 calls initially showed  
discordance between the method of the present invention and DNA sequencing. The 14  
discordant SNPs were re-sequenced and were found to correspond with the results obtained  
using the method of the present invention. Thus, following sample re-runs, the method of the  
30 present invention showed 100% accuracy for all calls made when compared to sequencing.

**[00125]** The ability of the method of the present invention to discriminate among the three  
possible calls (homozygous wild-type, heterozygote and homozygous mutant) is illustrated in  
Figure 4. The data clearly demonstrates that each of the call types is tightly clustered within  
an area that is well-defined and distinct from the other two call types. In other words, overlap  
35 between homozygotes and heterozygotes is not likely to occur. This large separation between  
signals generated for a homozygous wild-type, homozygous mutant or heterozygous sample  
greatly reduces the risk of miscalls, an extremely important requirement for any diagnostic

5 test. The variation in signal intensities between the different SNPs varies and is likely due to the different priming efficiencies of the ASPE primers and/or the number of incorporated biotin-dCTP nucleotides.

## **EXAMPLE 2: MULTIPLEX GENOTYPING OF THROMBOSIS**

[00126] Figure 5 depicts a flow chart overview of a methodology for genotyping of  
10 thrombosis using the present invention in conjunction with two distinct detection platforms, specifically the Luminex and Metrigenix microarray platforms. In the method illustrated in figure 5, six specific thrombosis-associated SNPs are tested 60. Genomic DNA is subjected to a multiplex amplification reaction 70 to prepare amplified regions of DNA containing the six specific thrombosis-associated SNPs. Tagged ASPE primers are then hybridized to the  
15 amplified regions, each tagged allele specific primer having a 3' portion complementary to a region of the amplified DNA, a 3' terminal nucleotide complementary to one allele of one of the mutation sites (wild type or mutant) mentioned above, and a 5' portion complementary to a probe sequence. Extension reactions are initiated, wherein a labelled extension product of the primer is synthesised when the 3' terminal nucleotide of the ASPE primer is  
20 complementary to a corresponding nucleotide in the target sequence; no extension product is synthesised when the terminal nucleotide of the primer is not complementary to the corresponding nucleotide in the target sequence. Labelled extension products are then captured on a universal array 90. Detection of extension products can be accomplished either using the Luminex xMAP platform 100, or using an MGX Flow Thru Chip 110. Software is  
25 used to interpret the data generated, and to identify the genotype of the individual being tested 120.

[00127] All six thrombosis related mutations selected from the group consisting of factor V Leiden G1691A, factor II G20210A (prothrombin), MTHFR C677T, MTHFR A1298C, Factor XIII val34leu and TFPI C536T, were tested for using a combination of the Luminex  
30 bead-based array and the Metrigenix Flow-Thru chips. Multiplex PCR amplification of a sample was conducted, as described above, in order to amplify the six polymorphic regions. Allele specific primer extension, as described further above, was also conducted. Extension products were labelled using biotin.

[00128] Figures 6 and 7 present the results of a single patient sample genotyped using the  
35 universal array format on the Luminex Bead-Based platform (figure 6) and the Metrigenix Flow-Thru Chip platform (figure 7). Biotinylated extension products in the sample tested were detected through the interaction of the biotin with streptavidin-phycoerythrin when the

5 Luminex bead-based platform was used. Biotinylated extension products in the sample tested were detected through the interaction of the biotin with streptavidin-horseradish peroxidase in the presence of luminal and hydrogen peroxide when the Metrigenix Flow-Thru Chip platform was used.

10 [00129] The results demonstrate the accuracy and specificity of the method of the present invention. All genotypes were correctly identified using both array platforms, and the allelic ratios which were determined using each platform were almost identical.

15 [00130] Although the invention has been described with reference to certain specific embodiments, various modifications thereof will be apparent to those skilled in the art without departing from the spirit and scope of the invention as outlined in the claims appended hereto.